A 15-Year Longitudinal Study of Halstead-Reitan Neuropsychological Test Performance

Merrill F. Elias, Michael A. Robbins, and Penelope K. Elias

Methods

Participants

Beginning in 1975, 107 community-dwelling individuals participated in a cross-sectional study of hypertension and performance on the Wechsler Adult Intelligence Scale and in this context were examined for the presence of hypertension-related target organ disease and other diseases (e.g., Schultz, Dineen, Elias, Pentz, & Wood, 1979). The present sample consists of a subset of these individuals (n = 53) who participated in this longitudinal study of neuropsychological test performance. Occupations ranged from blue-collar worker to professional. Means and standard deviations for age and education level were as follows: age, M = 44.7 years, SD = 12.5, range = 20–67; education, M = 14.9 years, SD = 2.0, range = 10–20 years.

Participants were evaluated four times, at 5- to 6-year intervals. At each examination a minimum of 15 blood pressure measurements and a complete medical history were taken. Each participant completed the Cornell Medical Index (Brodman, Ermann, & Wolff, 1956), and a careful record of diagnoses and treatments, based on patient records, physician report, and patient self-report was kept.

Individuals with any of the following histories or diagnoses were excluded from the sample: use of psychotropic drugs, alcohol abuse, diabetes, neurological disorder, senile dementia, brain trauma, physician-diagnosed mental illness, cerebral vascular disease (including transient ischemic attack), cardiovascular disease (including angina and myocardial infarction), renal disease, catastrophic chronic or acute disease. Thirty-six of the 53 subjects were tested at all four exams, 13 were tested at three exams, and 4 were tested at only two exams. Reasons for attrition between examinations 3 and 4 were: death (n = 3), inability to participate due to illness (n = 4), and refusal to participate (n = 1). Reasons for attrition between Exams 2 and 3 were: death (n = 2) and refusal to participate (n = 2).

A health status variable, to be used as a covariate (statistical control), was constructed by index-coding the participants as 0 (Continuously Healthy [CH]) or 1 (Not Continuously Healthy [NCH]). The CH group (n = 25) included only those individuals who were not diagnosed and treated for hypertension at any examination and continued to be free from all
diseases and symptoms resulting in exclusion from the study at baseline. The NCH group \((n = 28)\) had been diagnosed as hypertensive (systolic and/or diastolic blood pressure \(> 140\) mmHg and/or \(90\) mmHg, respectively) at Exam 1, or experienced the following illnesses during the course of the study: myocardial infarction \((n = 2);\) diabetes \((n = 1);\) family report of a diagnosis of Alzheimer’s disease \((n = 1);\) mild stroke \((n = 2);\) surgery for carotid artery disease \((n = 1);\) hypertension diagnosed after Exam 1 \((n = 1)\).

Procedures and Measures

Procedures for testing have been reported in detail elsewhere (e.g., Elias et al., 1990). All participants who did not have normal vision were asked to wear glasses, and testing was conducted by a neuropsychological examiner after screening for aphasia (none detected), and determining eye and hand dominance (unrelated to any of the findings reported).

The tests, widely used clinically and described elsewhere (Elias et al., 1990; Reitan & Wolfson, 1993), were given in the following order: Digit Symbol Substitution (DS), Tactual Performance Test (TPT) — time component (TPT-TIME), TPT — memory for forms (TPT-MEM), TPT — localization of forms from memory (TPT-LOC), Finger Tapping (FT) with dominant hand (FT-DH) and nondominant hand (FT-NDH), Trail Making Tests A and B (Trails A, Trails B), and the Category Test (CAT). TPT-TIME is derived by summing time to complete the TPT with the dominant hand alone, the nondominant hand alone, and with both hands. Seven of the scores can be used to calculate an overall, prorated AIR score (Russell et al., 1970), and relations among test scores can be compared in order to infer locus and extent of brain injury if AIR exceeds 1.55, the cutoff for young adults (Russell et al., 1970). The FT score used in the calculation of AIR is based on either FT-DH or FT-NDH, depending on the hand for which performance is lower (more impaired). Trails A is given prior to Trails B, but is not used in the calculation of AIR.

The AIR score is the mean of the “Rating Equivalent” (RE) scores for each test. Following Russell et al.’s (1970) formula, in order to calculate AIR the raw test scores were transformed to a 6-point RE score as follows: 0 = above average; 1 = average; 2 = mild impairment; 3 = moderate impairment; 4 = severe impairment; 5 = very severe impairment.

Longitudinal data for the CAT were not available for 8 participants because they did not complete the CAT at on at least two occasions. Thus, for all 53 participants, AIR was calculated excluding CAT; the 45 participants for whom longitudinal CAT data were available also received AIR scores which included their CAT score.

Statistical Analyses

A two-stage growth-curve method (Rogosa, Brandt, & Zimowski, 1982) was employed, allowing us to retain all individuals who completed at least two examinations. Neither equal numbers of participants at each longitudinal examination, nor equal time intervals between longitudinal examinations are required (Aldwin, Spiro, Levenson, & Bossé, 1989; Rogosa et al., 1982; Willett, 1988/1989).

Paraphrasing Aldwin et al. (1989), for Stage 1 of the analyses a linear model is fit to the neuropsychological data obtained for each individual using the method of least squares: \(Y_a = a_1 + b_1t_1 + e_1\). For each individual \(i\) at each time of measurement \(t_i\), \(Y\) is the observed test score, \(a\) is the intercept, and \(b\) is the raw regression coefficient for test scores regressed on time. The intercept \((a)\) gives each participant’s estimated score at entry into the study, and the slope \((b)\) gives each participant’s predicted change in the test scores per year since baseline (Exam 1). Linear, rather than nonlinear curves, were fit to the test scores because, as Aldwin et al. (1989) point out, the linear fit provides a good approximation of the quadratic fit when the number of examinations is limited.

Thus, for Stage 1 of the analyses, individual test scores were regressed on years from the point of entry into the study (Exam 1). The predicted slope and 0-intercept derived from Stage 1 for each individual on each test constituted the dependent variables for Stage 2 analyses. The predictor variable for the Stage 2 multiple regression analyses was participant’s age at baseline (in years); covariables included education (in years), gender, occupation, and the CH versus NCH vector. The quadratic \((\text{Age}^2)\) term did not enter into the model significantly in the presence of the age main effect (all \(p > .10\)) and was deleted from the model.

A weighted least-squares solution for slope values (Willett, 1988/1989, pp. 408-411) was used to allow values estimated with smaller standard errors (i.e., those estimated with more precision) at Stage 1 to be weighted more than those with larger standard errors.

RESULTS

Results of the Stage 1 analysis are shown in Table 1. The intercept and slope values are means \((M)\) of the predicted intercept and slope values obtained for each participant. Intercepts represent the means of predicted scores at Exam 1 (baseline). The mean slope values that differed from zero all indicated performance decline over time, for AIR scores (calculated with or without CAT), FT-DH, FT-NDH, DS, TPT-MEM, and TPT-LOC. Mean slope values did not differ significantly from zero for CAT, TPT-TIME, Trails A, and Trails B.

Using AIR (No-CAT) as an example, intercept and slope values may be interpreted as follows. The mean AIR value at the 0-intercept was .798. The mean AIR slope value of .025 indicates that over the course of the study (i.e., 15 years), performance declined an average of .025 points per year (0.0 to 5.0 scale). Extrapolating across the 15 years of the study, performance on the AIR declined an average of .375 points.

Stage 2 of the analysis was preceded by an evaluation of the impact of health status. The NCH (not continuously healthy) group exhibited greater performance decline over time (i.e., had larger slope values) than the CH group for AIR \((b = .0252, p < .01);\) FT-DH \((b = -.3274, p < .01);\) and FT-NDH \((b = -.3654, p < .01).\) The results were in the same direction for every other test score. Similarly, diastolic blood pressures, averaged over all examinations, were associated with greater decline \((p < .05)\) for both AIR scores, FT-NDH, DS, CAT, and Trails A and Trails B. However, both health covariables resulted in only minor and
Insignificant attenuation in magnitude of association between age and slope or age and intercept values. No significant interactions between age and health status were obtained (all $p > .10$). Data for the Stage 2 analyses presented below and in Table 2 include health status (CH vs NCH) as a covariate.

In Stage 2, multiple regression analyses were performed in which slope and intercept values for each test score were regressed on age (the predictor variable), education, occupation, gender, and health status (the covariates). Each regression coefficient was controlled for all others. Thus, the coefficients shown for the Stage 2 analysis represent the independent association of age with intercept and slope values. Where significant ($p < .05$) associations were observed, covariate-adjusted $R^2$ values ranged from .06 to .25, and for nonsignificant associations, they ranged from .00 to .05. However, the most important question was with respect to magnitude of associations between age and individual slope values (given by the regression coefficients in Table 2).

Increasing age at baseline (Exam 1) was associated with greater declines in performance over time for both AIR scores, FT–DH, FT–NDH, and DS. No statistically significant associations between age and individual slopes were observed for the other test scores ($ps > .05$). Even where associations reached statistical significance, the magnitude of change was small. For example, extrapolating from 1-year increments to 10-year increments of age at baseline, for every additional 10 years of age, performance on AIR declined by .165 points over the course of the study (15 years).

Age was unrelated to intercept (baseline) values, but was significantly (though modestly) related to slope (change over time) values for FT with both the dominant (FT–DH) and nondominant (FT–NDH) hands. Thus, although age made no difference to performance at Exam 1, the older individuals showed greater decline in FT over the next 15 years. On the other hand, for CAT and TPT–TIME, older participants performed more poorly at baseline (intercept), but did not decline at a greater rate than the younger participants (slope).

**DISCUSSION**

Both health covariates (CH vs NCH, diastolic blood pressures) were associated with poorer performance for AIR scores, FT–DH, FT–NDH, and DS. However, significant associations reached statistical significance

<table>
<thead>
<tr>
<th>Test Measure</th>
<th>Intercept Slope*</th>
<th>Mean Slope*</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIR (No-Category Test)</td>
<td>0.793</td>
<td>0.025**</td>
</tr>
<tr>
<td>(0-5 points)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIR (With Category Test)</td>
<td>0.893</td>
<td>0.022**</td>
</tr>
<tr>
<td>(0-5 points)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finger Tapping (DH)*</td>
<td>52.644</td>
<td>-0.264**</td>
</tr>
<tr>
<td>(taps/10 sec)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finger Tapping (NDH)*</td>
<td>47.728</td>
<td>-0.231**</td>
</tr>
<tr>
<td>(taps/10 sec)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digit Symbol</td>
<td>61.259</td>
<td>-0.256**</td>
</tr>
<tr>
<td>(No. correct in 1.5 min)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category Test</td>
<td>37.784</td>
<td>-0.073</td>
</tr>
<tr>
<td>(errors)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPT–Total Time</td>
<td>11.449</td>
<td>0.024</td>
</tr>
<tr>
<td>(minutes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPT–Memory</td>
<td>7.797</td>
<td>-0.036*</td>
</tr>
<tr>
<td>(No. correct of 10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPT–Localization</td>
<td>5.113</td>
<td>-0.087**</td>
</tr>
<tr>
<td>(No. correct of 10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trails A</td>
<td>24.325</td>
<td>0.007</td>
</tr>
<tr>
<td>(seconds)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trails B</td>
<td>59.027</td>
<td>0.005</td>
</tr>
</tbody>
</table>

*Significant at $p < .05$.
**Significant at $p < .01$.
*Significant at $p < .001$.

**Table 1. Results of the Stage 1 Analysis Showing the Mean of the Predicted Intercept and Slope Values**

**Table 2. Results of the Stage 2 Analysis Showing Regression Coefficients (β) Expressing the Relations Between Age and Estimated Initial Performance (Intercept) and Age and Change in Performance for Each Year of Participation (Slope)**

**Table 3. Results of the Stage 2 Analysis Showing Regression Coefficients (β) Expressing the Relations Between Age and Estimated Initial Performance (Intercept) and Age and Change in Performance for Each Year of Participation (Slope)**

*Significant at $p < .05$.
**Significant at $p < .01$.
***Significant at $p < .001$.
and one or more of the individual test measures. However, including these covariates in the regression model had a trivial, nonsignificant influence on relations between age and magnitude of change over time. This phenomenon has been reported in previous studies (Earles & Salthouse, 1995; Salthouse, Kauser, & Sauls, 1990), and may be related to the fact that cognitive functioning is influenced by so many biological and social psychological factors that controlling for health (other than for catastrophic or serious end-stage diseases when they are highly prevalent in the study sample), does not have an appreciable effect on relations between age and cognitive functioning (Elias, Elias, & Elias, 1990).

Despite the fact that each of the measures used in the present study has been associated with significant age-related decline in performance in cross-sectional studies (Elias et al., 1993), performance did not decline over time for CAT, TPT–TIME, Trails A and Trails B; and age was unrelated to longitudinal decline for these tests and for TPT–MEM and TPT–LOC.

Interestingly, the tests which make heavy demands on the ability to plan, organize, and execute responses according to a rule (executive functioning), i.e., the Category Test, Trails B (Boone, Miller, & Lesser, 1993), and the Tactual Performance Tests (Golden, Zillmer, & Spiers, 1992), showed meager and nonsignificant associations between age and change over time in cognitive performance. The tests that showed statistically significant, though modest, change over time were Finger Tapping and Digit Symbol. Finger Tapping, as administered in the HRB, is a primary motor skill (Fowler, Zillmer, & Newman, 1988), and Digit Symbol places disproportionately greater demands on speed of psychomotor performance than memory and executive functioning (Salthouse, 1978, 1990). In this regard, our results are in agreement with work by Salthouse (1990), which suggests that those psychomotor tests emphasizing speed of performance show the greatest age-related change.

Considering general, rather than specific, change in performance over time, a mean decline in AIR scores was seen for the individuals in our study (Table 1). Although the extent of decline was positively related to age at entry (Table 2), this relationship was clinically insignificant and was largely due to specific declines in FT and DS.

Clearly our participants were relatively well-educated, and thus our findings are limited with respect to generalizability to persons of lower educational levels. Thus, comprehensive normative data will require further studies across a broad educational range.

ACKNOWLEDGMENTS

This research was supported in part by National Institute on Aging research grant 5-R37-AG03055-13 to Merrill F. Elias, and in part by National Institute on Aging research grant 1-K01-AG00646-01 (Special Career Emphasis Award) to Penelope K. Elias.

Reprints and data with respect to the mean performance values at each examination and for dominant, nondominant, and both hands for the Tactual Performance Test may be obtained by corresponding with Dr. Merrill F. Elias, Department of Psychology, 5742 Little Hall, University of Maine, Orono, ME 04469-5742. E-mail: mfelias@maine.maine.edu

REFERENCES


Received January 27, 1995
Accepted June 12, 1996